

Inbreeding as Measured by Isonymy, Pedigrees, and Population Size in Törbel, Switzerland

WALTER S. ELLIS¹ AND WILLIAM T. STARMER²

INTRODUCTION

A shortcut method of measuring inbreeding, based on the hypothetical genetic relationship of married pairs who share the same surname, was formally presented by Crow and Mange [1]. This sharing of surnames, or isonymy, was first investigated by Darwin [2] and has provided a useful index of inbreeding trends in populations where marriage records are available [3, 4]. The hypothetical relationship between inbreeding and isonymy depends largely on the following assumptions: (1) all surnames in the populations are transmitted regularly and possible deviations (illegitimacy, adoption, change of name, etc.) are of negligible effect; (2) no one type of cousin mating is preferential or prohibited (random sex labeling of pedigrees); (3) neither sex is disproportionately represented among migrants into or out of the population; (4) the maximum mean inbreeding coefficient (F) is .25, because the method does not apply to marriages in which the partners are sibs; and most importantly, (5) isonymy necessarily reflects common ancestry (monophyletism).

In Kippel, Switzerland [5, 6], the isonymy method produced an inbreeding estimate that greatly exceeded that produced by the pedigree method. This was ascribed to a prevalence of isonymous mates within consanguinity classes, which was probably due to partible inheritance and patrilocality. The ecological setting of such a village results in a proliferation of relatives within the pool of potential mates, which yields a high level of random inbreeding. The number of patriline exceeded the number of matriline in the records of this village, biasing the estimates of inbreeding based on isonymy so that random isonymy was not equal to random inbreeding.

In this paper, we test the relationship between inbreeding estimated by isonymy and by pedigrees using records from Törbel, Switzerland. In addition, estimates of inbreeding based on population size provide data on the relationship between demographic parameters and the genetic structure of alpine villages.

Received June 12, 1977; revised January 2, 1978.

This paper was presented at the annual meeting of the Southwest and Rocky Mountain Division, American Association for the Advancement of Science, Tucson, Arizona, April 1976.

¹ Department of Anthropology, University of Arizona, Tucson, Arizona 85721.

² Department of Biology, Syracuse University, Syracuse, New York 13210.

© 1978 by the American Society of Human Genetics. All rights reserved.

TÖRBEL, SWITZERLAND

Törbel is located in the German-speaking part of the canton of Valais in the Vispental, a subsidiary Rhone River Valley, and lies at an altitude of 1,500 meters. The population of this village (approximately 600 in 1970) has remained isolated and homogeneous. Until the 1940s, 84.33% of all marriages were village endogamous. Moreover, in 300 years, no new paternal lines were introduced in the village, and all of the 12 currently resident name groups were represented by lineal ancestors in 1665. All of the present residents are German-speaking and Roman Catholic.

The pattern of village endogamy, in which, before World War II, approximately 80% of males and 90% of females took spouses from within the community, was promoted by the following factors: (1) partible inheritance where each child received an equal share of the parental estate; (2) denial of communal rights or citizenship to landowners from other villages or in-migrant male spouses unless they paid substantial sums for admission [7]; and (3) lack of significant emigration except during the late nineteenth century.

ORGANIZATION OF DATA

Our data was taken from a village genealogy compiled in the mid-nineteenth century by a local priest from records going back to the late seventeenth century. The genealogy records [8], which are organized into surname groups, were supplemented with whatever additional data could be found in civil and church records of births, marriages, and deaths.

Using a system called SELGEM (self-generating master) developed by the Smithsonian Institution, the Törbel master file was created, containing all the basic data on each individual. SELGEM is a generalized information processing system of COBOL programs that has retrieval, sorting, indexing, and report writing capabilities. The Törbel master file now formed the basis for a nuclear family file, from which the Törbel genealogy was created. All together, three FORTRAN programs were written and six SELGEM setups (combinations of COBOL programs) were employed in the process. The records of more than 5,500 individuals over more than 300 years were ordered into about 1,000 nuclear families.

PEDIGREE INBREEDING OVER TIME

Twenty-five year birth cohorts [9] based on the mean year of birth of the sibship were formed. Mean coefficients of inbreeding were derived from the computer program published by MacCluer et al. [10] and adapted to the Törbel data by Larry Manire of the University of Arizona Computer Center Research Support Section. The distribution of these coefficients for the period 1700–1974 is given in the left column of table 1. Low values of \bar{F} during the first few generations are an artifact due to incomplete pedigrees, but the estimates seem reliable after 1800, 6 generations after the parish records began. Thereafter \bar{F} increased to a peak of .007139 among the offspring born between 1925 and 1950, and then declined (table 1).

When inbreeding is broken down into close (second cousin or closer) and remote (beyond second cousin to a limit of 3.8146×10^{-6} which is the inbreeding coefficient of the offspring of two eighth cousins) components, we found that most inbreeding was

TABLE 1
MEAN COEFFICIENTS OF PEDIGREE INBREEDING AND RANDOM KINSHIP BY 25-YEAR PERIODS

Pedigree \bar{F} *	No.	Cohort	Random Kinship†	No.
...	62	1700	...	62
...	56	1725	...	56
.001689	37	1750	.003438	200
.004357	52	1775	.002656	200
.006165	67	1800	.009902	200
.002941	85	1825	.006309	200
.006073	87	1850	.005825	100
.006830	94	1875	.007225	60
.006350	110	1900	.006830	60
.007139	109	1925	.008540	100
.005829	190	1950	.007942	90

* Mean: .00496; total no.: 949.

† Mean: .00561; total no.: 1,328.

remote. Figure 1 presents these results as a percentage of families having remote, close, and no inbreeding from 1750 to 1950.

Simulation of Pedigree Inbreeding

A modification of the MacCluer et al. [10] program was used to generate random pairs of parental sibships from all possible pairs of siblings within 25-year periods. The

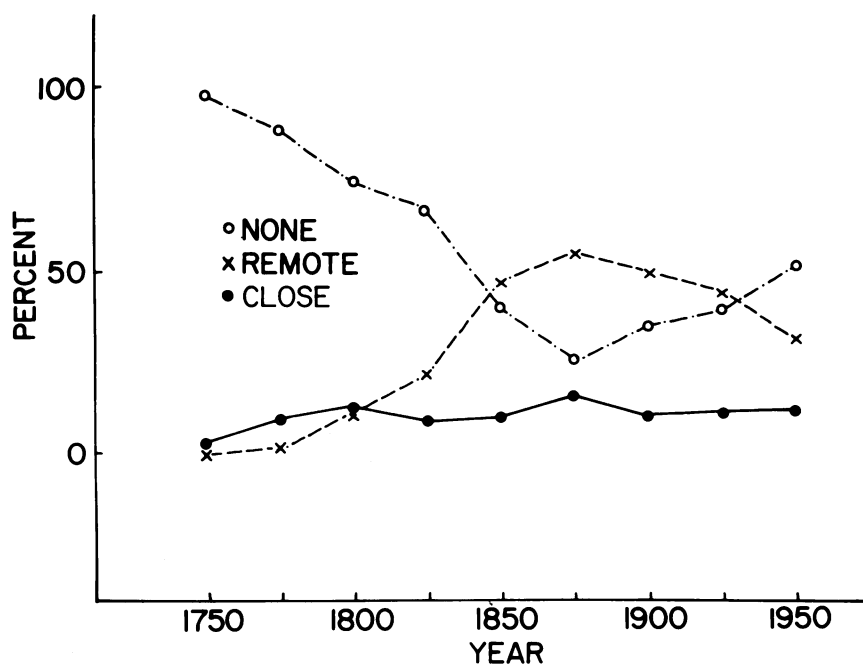


FIG. 1. — Percentage of families with close, remote, and no inbreeding by 25-year periods.

generation of pairs was terminated when the mean inbreeding coefficient for all pairs reached an approximate stable value. Sibship and ancestor-descendant pairings were not permitted. The mean inbreeding coefficient for random pairs was then compared to that based on actual pedigrees (right column of table 1). Inbreeding from simulation and pedigrees have similar means (.0056 and .0050). Mating is random in Törbel, at least for genealogical relationship within the constraints imposed by the simulation procedure.

A similar pattern was observed in the nearby Saas valley studied by Hussels [9]. Since the Saas valley was more isolated and poorer than Törbel, pedigree inbreeding tended to be slightly higher through time. It decreased earlier reflecting the buildup of the Saas valley for tourism around the turn of the century. As in Törbel, the simulation of random kinship for Saas produced almost identical means with pedigree inbreeding. At least in these two settings, random mating for genealogical relationship was occurring.

ESTIMATES OF INBREEDING USING POPULATION SIZE

Jacquard [11] has shown how the course of change of the mean inbreeding coefficient of a population is determined by the numbers of males and females in the population. His formula ([11], p. 498)* is equivalent to that of Crow and Kimura ([12], equation 3.11.1). If the value of the inbreeding coefficient for the population size is known, this equation can be used to calculate the value of inbreeding due to population size for any generation. Therefore, the Jacquard formula was used with the varying assumptions of population size given in table 2. The estimates were based on detailed information on in-marrying females, out-marrying males, and migrating families, which fortunately, the Swiss record. The offspring of female out-migrants may, however, be underrepresented, since only the records of males (Burghers) come back to the village for inclusion into the civil records.

TABLE 2
ESTIMATES OF POPULATION SIZE IN TÖRBEL BY 25-YEAR PERIODS

Year	No. Married Pairs	No. Married Pairs Who Reproduced	Corrected for Mean and Variance of Progeny No.*
1750	112	103	100
1775	74	73	90
1800	102	71	71
1825	128	115	88
1850	166	142	112
1875	166	144	105
1900	184	154	122
1925	212	164	155
1950	192	184	141

NOTE. — Estimates were made using SPSS procedures and original FORTRAN programs on the Törbel master file and nuclear family file.

* See Crow and Kimura [12], formula 7.62.17. Numbers are adjusted for effective no. of migrants.

* Here the formula is correct, but it is incorrect on page 163.

Figure 2 illustrates the results of the four calculations and shows that only f_4 resembles estimates of inbreeding from pedigrees, f_5 . Estimate f_4 uses an adaptation of formula 6.6.2 from Crow and Kimura [12] for autozygosity corrected for migration:

$$f_t = \left[\frac{1}{2N_{t-2}} + \left(1 - \frac{1}{N_{t-2}} \right) f_{t-1} + \frac{1}{2N_{t-2}} f_{t-2} \right] (1 - M_{t-2})^2.$$

Here M_{t+1} is equal to the number of mating pairs (P) in generation $t + 1$, minus the number of reproducing progeny (Q) from generation t , divided by (P). Thus, $(P_{t+1} - Q_t)/P_{t+1} = M_{t+1}$.

A paired sample t test with 7 degrees of freedom, comparing estimates f_4 and f_5 , yielded a t value of -1.20 with probability .265. We therefore accept the null hypothesis that the mean estimates of total inbreeding based on the estimate of population size with a correction for out-migration and on the pedigree calculation, do not differ. All other curves (f_1 , f_2 , and f_3) differ significantly from the pedigree estimate with $P < .1\%$.

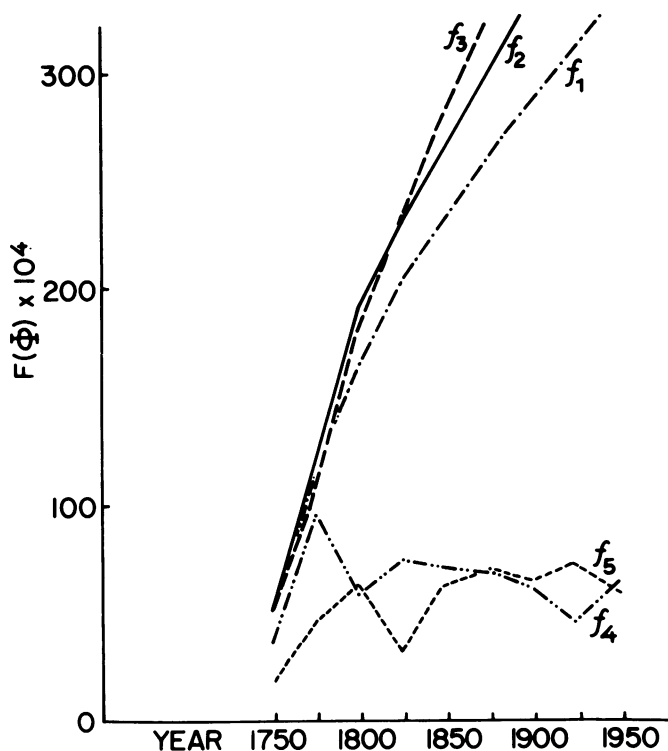


FIG. 2. — Coefficients of random kinship based on four different estimates of population size, 25-year periods. f_1 = no. of families; f_2 = no. of reproducing families; f_3 = no. using \bar{K} and $V_{\bar{K}}$; f_4 = f_1 adjusted for migration; and f_5 = pedigree inbreeding.

TABLE 3
FIVE METHODS OF CALCULATING RANDOM ISONYMY

Calculation*	Pt^r	Note
$\Sigma q_m q_f$1251	Proportion of certain name in males \times proportion of certain name in females
Σq_m^21240	Proportion of certain name in males
Σq_f^21345	Proportion of certain name in females
Σq_{m+f}^21272	Proportion of certain name in the population
$I_e = \frac{N}{N-1} (\Sigma q_i^2 - \frac{1}{N})$1206	Random isonymy from a unique sample. N = size of sample; Σq_i^2 = proportion of persons with the i th surname

* Summation is over all surnames.

THE ISONYMY METHOD

Of 195 parent pairs in the living population, 20 were isonymous (10.26%). Expected or random isonymy was calculated five different ways [13] (table 3), but little variation in Pt^r resulted.

Originally, Crow and Mange [1] calculated random inbreeding (F_{ST}) from $Pt^r/4$ where Pt^r was equivalent to $\Sigma q_m q_f$. Nonrandom inbreeding (F_{IS}) was calculated according to the formula $(Pt - Pt^r)/4(1 - Pt^r)$, where Pt is observed isonymy. Total inbreeding (F_{IT}) was then $F_{IS} + (1 - F_{IS})F_{ST}$. Yasuda and Furusho [13], noting that $F_{ST} = Pt^r/4$ and $F_{IT} = Pt/4$, derived nonrandom inbreeding (F_{IS}) from $(F_{IT} - F_{ST})/(1 - F_{ST})$. In Yasuda and Furusho's formulation, Pt and Pt^r are calculated directly from the data. These are then used to compute F_{IT} , F_{IS} , and F_{ST} . Crow [14] argues for the superiority of the Crow-Mange procedure over the Yasuda-Furusho method, but we employ both in this analysis.

Table 4 illustrates this comparison. Since there is little differentiation evident between results, the isonymy method can be said to estimate inbreeding in the living population of Törbel as follows: $F_{IS} = -.006$; $F_{ST} = .031$; and $F_{IT} = .025$. Here F_{IT} is five times the value calculated from pedigrees and the simulation runs ($F = .005$).

TABLE 4
CALCULATION OF INBREEDING BY ISONYMY METHOD: COMPARISON OF TWO PROCEDURES

	$q_m q_f$	q_m^2	q_f^2	q_{m+f}^2	$I_e(\text{pair})$
$F_{ST}(\text{random})$:					
Crow and Mange [1]0313	.0310	.0336	.0318	.0313
Yasuda and Furusho [13]0313	.0310	.0336	.0318	.0313
$F_{IS}(\text{nonrandom})$:					
Crow and Mange [1]	-.0064	-.0061	-.0092	-.0072	-.0064
Yasuda and Furusho [13]	-.0059	-.0056	-.0083	-.0064	-.0059
$F_{IT}(\text{total})$:					
Crow and Mange [1]0251	.0251	.0247	.0250	.0251
Yasuda and Furusho [13]0256	.0256	.0256	.0256	.0256

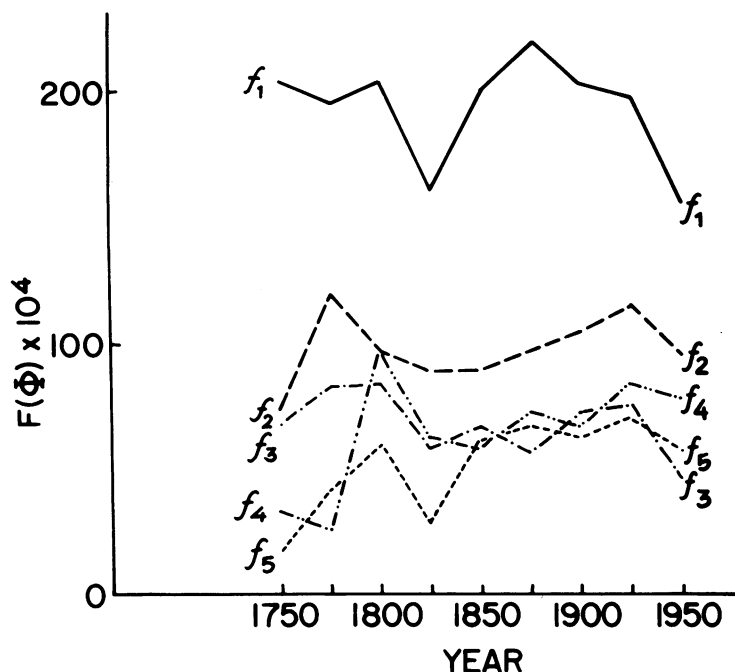


FIG. 3. — Inbreeding calculated from isonymy and pedigrees by 25-year periods. f_1 = inbreeding from isonymy; $f_2 = f_1$ corrected for polyphyletism; $f_3 = f_1$ corrected for polyphyletism and out-marriage; f_4 = simulation of random kinship; and f_5 = pedigree inbreeding.

Inbreeding was then calculated for 25-year marriage cohort periods using the method of Crow and Mange [1]. The result appears as f_1 in figure 3. According to this estimate, inbreeding peaks between 1875 and 1899 and then falls off. The relationship between inbreeding calculated from isonymy and inbreeding calculated from pedigrees is still unclear. A paired sample t test yields $t = 17.89$, which is larger than the .1% value of t (5.408) with 7 degrees of freedom; we, therefore, reject the null hypothesis at the .1% level of significance, concluding that the two methods of measuring inbreeding differ in the mean values they estimate.

Nonrandom Sex Labeling of Pedigrees

When estimating inbreeding from isonymy, we assume that all types of cousin marriages occur at random. Then in theory, the proportion of marriages between persons of the same family name should always reflect the same amount of inbreeding in the population, there always being one-fourth of the cousin marriages of each type between persons of the same family names. Hussels [9] and Morton and Hussels [15] regressed the probability of isonymy as a binary variable (0,1) on the inbreeding coefficient. The null hypothesis of random sex labeling of pedigrees is satisfied if the regression coefficient is equal to 4. In Saas, the regression coefficient was equal to 4.58 ± 0.72 [9]. A sample of 1,560 sibships with at least one case of retinitis pigmentosa, hemophilia (A,B), or myotonic dystrophy, ascertained through the Institut Univer-

TABLE 5
COEFFICIENTS OF REGRESSION OF ISONYMY ON MEAN PEDIGREE
INBREEDING COEFFICIENT BY 25-YEAR PERIODS

Year	$\beta \pm SE$
1775	3.99 ± 2.50
1800	6.58 ± 1.99
1825	$12.48 \pm 3.93^*$
1850	7.20 ± 2.74
1875	$9.61 \pm 2.60^*$
1900	3.87 ± 1.85
1925	$9.76 \pm 2.26^*$
1950	6.06 ± 1.77
Total	6.56 ± 0.76

* Significantly different from $\beta = 4$ ($P < .05$).

sitaire de Génétique Médicale, Geneva [15], produced a regression coefficient equal to 4.42 ± 0.57 . In Törbel, not all values are in good agreement with the value of 4 expected under the assumption of random sex-labeling of pedigrees, but the overall value is (see table 5).

The possibility remains that for remote consanguinity there is a significant excess of lines of descent through males. It is likely that a tendency toward patrilineal residence and the patrilineal nature of Swiss records lead to incomplete ascertainment of long lines of descent through females. The Törbel sample shows stronger preference for patrilineal chains of relationship than the Saas sample, but the difference is not significant. For a similar area (the Parma Valley), Cavalli-Sforza and Bodmer [16] demonstrated that pedigrees through males are more common (or, alternatively, that matrilineal chains are harder to trace).

Correction for Polyphyletism

Polyphyletism of names occurs when the same name derives from different sources, making two seemingly identical names nonisonymous. Cavalli-Sforza and Bodmer [16] analyzed possible sources of error in carrying over conclusions from surnames to genes, and Morton and Hussels [15] focused on the problem of polyphyletism.

Between 1950 and 1974, there were 92 endogamous marriages in Törbel. Of the 12 resident name groups, only three were monophyletic. A century earlier, five of 13 name groups were monophyletic. Monophyletic surnames are never even 50% of the total. The significance of this difference is shown quite clearly in inbreeding calculations. If all 219 living sibships had had different ancestors, then $F_{ST} = 1/(4 \times 219) = .00114$, which is less than one-tenth of the value obtained from the previous assumption of monophyletism. Correcting for polyphyletism and for out-marriage (see below) yields an F_{ST} of .0047 which is consistent with other estimates.

Using a program called SURNAME,* written by the staff of the University of Arizona

* This program can be obtained from W. S. Ellis.

Research Support Section, lines of married pairs in Törbel were traced back to their initial fathers. Inbreeding was recalculated for those with identical names who were actually descended through males from a single ancestor. The resulting inbreeding coefficients were compared to F uncorrected for polyphyletism and to F calculated from pedigrees. The inbreeding coefficient declines an average of 51.6% after the polyphyletism correction. However, mean values of inbreeding calculated from pedigrees and inbreeding calculated from isonymy with correction for polyphyletism still differ significantly: $t = -9.23$, $P < .1\%$. The relationship between the results of these methods is depicted in figure 3 (f_2 and f_5).

Correction for Out-marriage

A correction to the isonymy method for out-marriage was made because there is no probability of isonymy for unique names (unless surnames are present regionally, which is the equivalent of a marriage to an unrelated individual), and because their frequency varied considerably through time, reaching a peak after World War II. Since only endogamous marriages can be isonymous (with only few exceptions), isonymy was again recalculated, as follows: $F_{ST} = \Sigma p_i q_i / 4$ became $\Sigma p_i q_i / 4$ multiplied by the proportion of endogamous marriages in a particular time period. Similarly,

$$F_{IS} = \frac{P - \Sigma P_i q_i}{4(1 - \Sigma P_i q_i)} \text{ became } \frac{P - [\Sigma p_i q_i (\text{proportion endogamous})]}{4[1 - \Sigma p_i q_i (\text{proportion endogamous})]}$$

This calculation yielded f_3 of figure 3. Paired sample $t = 2.18$ for 7 degrees of freedom yielded a two-sided probability level of 6.1%. This is larger than the 5% level of t (2.37) with 7 degrees of freedom, and so we accept the null hypothesis at the 5% level of significance.

This resulting correction of isonymy together with the correction for polyphyletic lines yielded a mean estimate of inbreeding that does not differ significantly from that produced using pedigrees. It is of some significance here that the effect of exogamy was significant in a setting where migration itself was minimal. Therefore, the estimate of inbreeding from isonymy in nonisolated populations becomes a tenuous proposition, especially if the researcher is looking beyond simple relative trends.

Tests of the Use of Arbitrary Time Periods

In theory, the proportion of marriages between persons with the same family name should always reflect the same amount of inbreeding in the population, since there is always one-fourth of the cousin marriages of each type between persons of the same family names. This does not seem to be true in Törbel, since inbreeding estimated from isonymy greatly exceeds inbreeding estimated from pedigrees. It is possible that this discrepancy is a result of the arbitrary division of historical records and that some other division would perhaps not produce the same results. For instance, Küchemann et al. [17], in their work on Oxfordshire villages, and Roberts and Rawling [4], in their study of Northumberland parishes, employed a single set of interval divisions based largely on the number of observed occurrences of marriages in each period.

To discover how far such arbitrary interval divisions may have affected the results of isonymy calculations for Törbel, two further analyses were made. The marriage registers were re-examined, using both 30- and 50-year periods in the eighteenth and nineteenth centuries. In general the results of these analyses show considerable similarity among intervals, with the longer time period smoothing out the fluctuations that appear for the shorter time periods. As Roberts and Rawlings state, “too much reliance cannot be placed upon a single figure for a single period calculated by a single method” [4]. They go on, however, to state their belief that trends in isonymy reflect biological and micro-evolutionary forces at work. The comparison of inbreeding trends revealed by the isonymy method and by the pedigree method in Törbel (fig. 3) allows us to concur with this statement. However, the isonymy method clearly does not quantify these forces very well.

SUMMARY

Törbel provides an interesting test case for the study of the relationship between inbreeding measured by pedigrees and inbreeding measured by isonymy.

At the start of this investigation, we were aware that isonymy could introduce biases into the calculation of the inbreeding coefficient in either direction. However, it was expected that in Switzerland, inbreeding from isonymy would be an overestimate due to patrilocal residence and polyphyletic names. One way of dealing with this problem [13] was not to be concerned with the absolute value of inbreeding but only in the difference between estimates. Any bias introduced in the estimate itself disappears in such comparisons, so that a trend of inbreeding can be ascertained correctly. However, it was considered equally important to subject several populations to both a complete pedigree analysis and an isonymic analysis to determine the relationship between estimates of inbreeding. Despite the fact that several authors (Swedlund [18], for example) warned users of isonymy to exercise caution, the careless application of isonymy still persists.

In the present study, estimates of inbreeding from isonymy were brought into line with other methods based on pedigree analysis and population size. However, it was possible to do this only in Törbel where pedigree depth was extensive and relatively complete. Similar corrections are possible only when the distribution of mono- and polyphyletic names is known and when migration data are reliable. If the trouble is taken to make these corrections, the same time and effort might as well be spent in pedigree analysis (when fairly complete ascertainment is possible) to achieve the same end result.

ACKNOWLEDGMENTS

The authors wish to thank David Glenn Smith and Gabriel Lasker for their helpful comments and suggestions. Jean MacCluer was kind enough to provide us with a program for computing pedigree inbreeding for which we are especially grateful. Thanks are due to Robert Netting for kindly providing the Törbel data file to us.

REFERENCES

1. CROW JF, MANGE AP: Measurements of inbreeding from the frequency of marriages between persons of the same surname. *Eugen Q* 12:199–203, 1965
2. DARWIN GH: Marriages between first cousins in England and their effects. *J Stat Soc* 38:153–184, 1875
3. LASKER GW, CHIARELLI B, MASALI M, FEDELE F, KAPLAN BA: Degree of human genetic isolation measured by isonymy and marital distances in two communities in an Italian alpine valley. *Hum Biol* 44:351–360, 1972
4. ROBERTS DF, RAWLING CP: Secular trends in genetic structure: an isonymic analysis of Northumberland parish records. *Ann Hum Biol* 1:392–410, 1974
5. FRIEDL J, ELLIS WS: Inbreeding, isonymy, and isolation in a Swiss community. *Hum Biol* 46:699–712, 1974
6. ELLIS WS, FRIEDL J: Inbreeding as measured by isonymy and by pedigrees in Kippel, Switzerland. *Soc Biol* 23:158–167, 1976
7. NETTING RM: What alpine peasants have in common: observations on communal tenure in a Swiss village. *Hum Ecol* 4:135–146, 1976
8. KNODEL J, ESPENSHADE T: Genealogical studies as demographic data, in *Radhakamal Mukerjee Commemoration Volume*, edited by HUSIAN IA, Tunbridge Wells, Kent, Abacus Press, 1973
9. HUSSELS I: Genetic structure of Saas, a Swiss isolate. *Hum Biol* 41:469–479, 1969
10. MACCLUER JW, GRIFFITH R, SING CF, SCHULL WJ: Some genetic programs to supplement self-instruction in FORTRAN. *Am J Hum Genet* 19:189–221, 1967
11. JACQUARD A: *The Genetic Structure of Populations*. New York, Springer-Verlag, 1974
12. CROW JF, KIMURA M: *An Introduction to Population Genetics Theory*. New York, Harper and Row, 1971
13. YASUDA N, FURUSHO T: Random and nonrandom inbreeding revealed from isonymy study. Small cities of Japan. *Am J Hum Genet* 23:303–316, 1971
14. CROW JF: The estimation of inbreeding from isonymy. Madison, Univ. Wisconsin, paper no. 1607. Unpublished manuscript
15. MORTON NE, HUSSELS I: Demography of inbreeding in Switzerland. *Hum Biol* 42:65–78, 1970
16. CAVALLI-SFORZA LL, BODMER WF: *The Genetics of Human Populations*. San Francisco, W.H. Freeman, 1971
17. KÜCHEMANN CF, BOYCE AJ, HARRISON GA: A demographic and genetic study of a group of Oxfordshire villages. *Hum Biol* 39:251–276, 1967
18. SWEDLUND AC: Isonymy: estimating inbreeding from social data. *Bull Eugen Soc* 7:67–73, 1975